Linear Algebra

and its applications

FOURTH EDITION

David C. Lay

6 Orthogonality and Least Squares

INTRODUCTORY EXAMPLE

The North American Datum and GPS Navigation

Imagine starting a massive project that you estimate will take ten years and require the efforts of scores of people to construct and solve a 1,800,000 by 900,000 system of linear equations. That is exactly what the National Geodetic Survey did in 1974, when it set out to update the North American Datum (NAD)—a network of 268,000 precisely located reference points that span the entire North American continent, together with Greenland, Hawaii, the Virgin Islands, Puerto Rico, and other Caribbean islands.

The recorded latitudes and longitudes in the NAD must be determined to within a few centimeters because they form the basis for all surveys, maps, legal property boundaries, and layouts of civil engineering projects such as highways and public utility lines. However, more than 200,000 new points had been added to the datum since the last adjustment in 1927, and errors had gradually accumulated over the years, due to imprecise measurements and shifts in the earth's crust. Data gathering for the NAD readjustment was completed in 1983.

The system of equations for the NAD had no solution in the ordinary sense, but rather had a *least-squares solution*, which assigned latitudes and longitudes to the reference points in a way that corresponded best to the 1.8 million observations. The least-squares solution was found in 1986 by solving a related system of so-called *normal equations*, which involved 928,735 equations in 928,735 variables.[1]

More recently, knowledge of reference points on the ground has become crucial for accurately determining the locations of satellites in the satellite-based *Global Positioning System (GPS)*. A GPS satellite calculates its position relative to the earth by measuring the time it takes for signals to arrive from three ground transmitters. To do this, the satellites use precise atomic clocks that have been synchronized with ground stations (whose locations are known accurately because of the NAD).

The *Global Positioning System* is used both for determining the locations of new reference points on the ground and for finding a user's position on the ground relative to established maps. When a car driver (or a mountain climber) turns on a GPS receiver, the receiver measures the relative arrival times of signals from at least three satellites. This information, together with the transmitted data about the satellites' locations and message times, is used to adjust the GPS receiver's time and to determine its approximate location on the earth. Given information from a fourth satellite, the GPS receiver can even establish its approximate altitude.

[1] A mathematical discussion of the solution strategy (along with details of the entire NAD project) appears in *North American Datum of 1983*, Charles R. Schwarz (ed.), National Geodetic Survey, National Oceanic and Atmospheric Administration (NOAA) Professional Paper NOS 2, 1989.

Both the NAD and GPS problems are solved by finding a vector that "approximately satisfies" an inconsistent system of equations. A careful explanation of this apparent contradiction will require ideas developed in the first five sections of this chapter.

WEB

In order to find an approximate solution to an inconsistent system of equations that has no actual solution, a well-defined notion of nearness is needed. Section 6.1 introduces the concepts of distance and orthogonality in a vector space. Sections 6.2 and 6.3 show how orthogonality can be used to identify the point within a subspace W that is nearest to a point $\mathbf{y}$ lying outside of W. By taking W to be the column space of a matrix, Section 6.5 develops a method for producing approximate ("least-squares") solutions for inconsistent linear systems, such as the system solved for the NAD report.

Section 6.4 provides another opportunity to see orthogonal projections at work, creating a matrix factorization widely used in numerical linear algebra. The remaining sections examine some of the many least-squares problems that arise in applications, including those in vector spaces more general than $\mathbb{R}^n$.

6.1 INNER PRODUCT, LENGTH, AND ORTHOGONALITY

Geometric concepts of length, distance, and perpendicularity, which are well known for $\mathbb{R}^2$ and $\mathbb{R}^3$, are defined here for $\mathbb{R}^n$. These concepts provide powerful geometric tools for solving many applied problems, including the least-squares problems mentioned above. All three notions are defined in terms of the inner product of two vectors.

The Inner Product

If $\mathbf{u}$ and $\mathbf{v}$ are vectors in $\mathbb{R}^n$, then we regard $\mathbf{u}$ and $\mathbf{v}$ as $n \times 1$ matrices. The transpose $\mathbf{u}^T$ is a $1 \times n$ matrix, and the matrix product $\mathbf{u}^T\mathbf{v}$ is a 1×1 matrix, which we write as a single real number (a scalar) without brackets. The number $\mathbf{u}^T\mathbf{v}$ is called the **inner product** of $\mathbf{u}$ and $\mathbf{v}$, and often it is written as $\mathbf{u}\cdot\mathbf{v}$. This inner product, mentioned in the exercises for Section 2.1, is also referred to as a **dot product**. If

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \quad \text{and} \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

then the inner product of $\mathbf{u}$ and $\mathbf{v}$ is

$$\begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1v_1 + u_2v_2 + \cdots + u_nv_n$$

EXAMPLE 1 Compute $\mathbf{u}\cdot\mathbf{v}$ and $\mathbf{v}\cdot\mathbf{u}$ for $\mathbf{u} = \begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 3 \\ 2 \\ -3 \end{bmatrix}$.

SOLUTION

$$\mathbf{u}\cdot\mathbf{v} = \mathbf{u}^T\mathbf{v} = [\,2 \quad -5 \quad -1\,]\begin{bmatrix} 3 \\ 2 \\ -3 \end{bmatrix} = (2)(3) + (-5)(2) + (-1)(-3) = -1$$

$$\mathbf{v}\cdot\mathbf{u} = \mathbf{v}^T\mathbf{u} = [\,3 \quad 2 \quad -3\,]\begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix} = (3)(2) + (2)(-5) + (-3)(-1) = -1 \quad \blacksquare$$

It is clear from the calculations in Example 1 why $\mathbf{u}\cdot\mathbf{v} = \mathbf{v}\cdot\mathbf{u}$. This commutativity of the inner product holds in general. The following properties of the inner product are easily deduced from properties of the transpose operation in Section 2.1. (See Exercises 21 and 22 at the end of this section.)

THEOREM 1

Let $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ be vectors in $\mathbb{R}^n$, and let c be a scalar. Then

a. $\mathbf{u}\cdot\mathbf{v} = \mathbf{v}\cdot\mathbf{u}$
b. $(\mathbf{u} + \mathbf{v})\cdot\mathbf{w} = \mathbf{u}\cdot\mathbf{w} + \mathbf{v}\cdot\mathbf{w}$
c. $(c\mathbf{u})\cdot\mathbf{v} = c(\mathbf{u}\cdot\mathbf{v}) = \mathbf{u}\cdot(c\mathbf{v})$
d. $\mathbf{u}\cdot\mathbf{u} \geq 0$, and $\mathbf{u}\cdot\mathbf{u} = 0$ if and only if $\mathbf{u} = \mathbf{0}$

Properties (b) and (c) can be combined several times to produce the following useful rule:

$$(c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p)\cdot\mathbf{w} = c_1(\mathbf{u}_1\cdot\mathbf{w}) + \cdots + c_p(\mathbf{u}_p\cdot\mathbf{w})$$

The Length of a Vector

If $\mathbf{v}$ is in $\mathbb{R}^n$, with entries $v_1, \ldots, v_n$, then the square root of $\mathbf{v}\cdot\mathbf{v}$ is defined because $\mathbf{v}\cdot\mathbf{v}$ is nonnegative.

DEFINITION

The **length** (or **norm**) of $\mathbf{v}$ is the nonnegative scalar $\|\mathbf{v}\|$ defined by

$$\|\mathbf{v}\| = \sqrt{\mathbf{v}\cdot\mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}, \quad \text{and} \quad \|\mathbf{v}\|^2 = \mathbf{v}\cdot\mathbf{v}$$

Suppose $\mathbf{v}$ is in $\mathbb{R}^2$, say, $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$. If we identify $\mathbf{v}$ with a geometric point in the plane, as usual, then $\|\mathbf{v}\|$ coincides with the standard notion of the length of the line segment from the origin to $\mathbf{v}$. This follows from the Pythagorean Theorem applied to a triangle such as the one in Fig. 1.

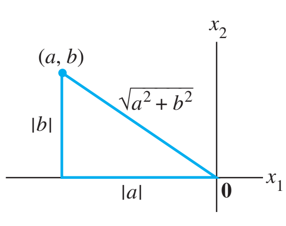


FIGURE 1
Interpretation of $\|\mathbf{v}\|$ as length.

A similar calculation with the diagonal of a rectangular box shows that the definition of length of a vector $\mathbf{v}$ in $\mathbb{R}^3$ coincides with the usual notion of length.

For any scalar c, the length of $c\mathbf{v}$ is $|c|$ times the length of $\mathbf{v}$. That is,

$$\|c\mathbf{v}\| = |c|\|\mathbf{v}\|$$

(To see this, compute $\|c\mathbf{v}\|^2 = (c\mathbf{v})\cdot(c\mathbf{v}) = c^2\mathbf{v}\cdot\mathbf{v} = c^2\|\mathbf{v}\|^2$ and take square roots.)

A vector whose length is 1 is called a **unit vector**. If we *divide* a nonzero vector $\mathbf{v}$ by its length—that is, multiply by $1/\|\mathbf{v}\|$—we obtain a unit vector $\mathbf{u}$ because the length of $\mathbf{u}$ is $(1/\|\mathbf{v}\|)\|\mathbf{v}\|$. The process of creating $\mathbf{u}$ from $\mathbf{v}$ is sometimes called **normalizing** $\mathbf{v}$, and we say that $\mathbf{u}$ is *in the same direction* as $\mathbf{v}$.

Several examples that follow use the space-saving notation for (column) vectors.

EXAMPLE 2 Let $\mathbf{v} = (1, -2, 2, 0)$. Find a unit vector $\mathbf{u}$ in the same direction as $\mathbf{v}$.

SOLUTION First, compute the length of $\mathbf{v}$:

$$\begin{aligned}\|\mathbf{v}\|^2 &= \mathbf{v}\cdot\mathbf{v} = (1)^2 + (-2)^2 + (2)^2 + (0)^2 = 9\\ \|\mathbf{v}\| &= \sqrt{9} = 3\end{aligned}$$

Then, multiply $\mathbf{v}$ by $1/\|\mathbf{v}\|$ to obtain

$$\mathbf{u} = \frac{1}{\|\mathbf{v}\|}\mathbf{v} = \frac{1}{3}\mathbf{v} = \frac{1}{3}\begin{bmatrix}1\\-2\\2\\0\end{bmatrix} = \begin{bmatrix}1/3\\-2/3\\2/3\\0\end{bmatrix}$$

To check that $\|\mathbf{u}\| = 1$, it suffices to show that $\|\mathbf{u}\|^2 = 1$.

$$\begin{aligned}\|\mathbf{u}\|^2 &= \mathbf{u}\cdot\mathbf{u} = \left(\tfrac{1}{3}\right)^2 + \left(-\tfrac{2}{3}\right)^2 + \left(\tfrac{2}{3}\right)^2 + (0)^2\\ &= \tfrac{1}{9} + \tfrac{4}{9} + \tfrac{4}{9} + 0 = 1\end{aligned}$$

■

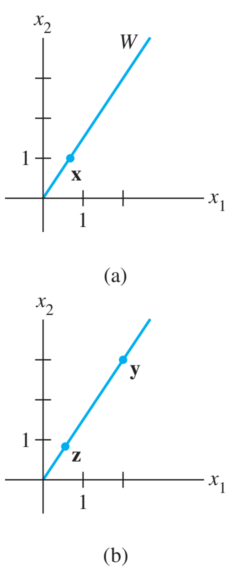


FIGURE 2
Normalizing a vector to produce a unit vector.

EXAMPLE 3 Let W be the subspace of $\mathbb{R}^2$ spanned by $\mathbf{x} = (\frac{2}{3}, 1)$. Find a unit vector $\mathbf{z}$ that is a basis for W.

SOLUTION W consists of all multiples of $\mathbf{x}$, as in Fig. 2(a). Any nonzero vector in W is a basis for W. To simplify the calculation, "scale" $\mathbf{x}$ to eliminate fractions. That is, multiply $\mathbf{x}$ by 3 to get

$$\mathbf{y} = \begin{bmatrix}2\\3\end{bmatrix}$$

Now compute $\|\mathbf{y}\|^2 = 2^2 + 3^2 = 13$, $\|\mathbf{y}\| = \sqrt{13}$, and normalize $\mathbf{y}$ to get

$$\mathbf{z} = \frac{1}{\sqrt{13}}\begin{bmatrix}2\\3\end{bmatrix} = \begin{bmatrix}2/\sqrt{13}\\3/\sqrt{13}\end{bmatrix}$$

See Fig. 2(b). Another unit vector is $(-2/\sqrt{13}, -3/\sqrt{13})$. ■

Distance in $\mathbb{R}^n$

We are ready now to describe how close one vector is to another. Recall that if a and b are real numbers, the distance on the number line between a and b is the number $|a - b|$. Two examples are shown in Fig. 3. This definition of distance in $\mathbb{R}$ has a direct analogue in $\mathbb{R}^n$.

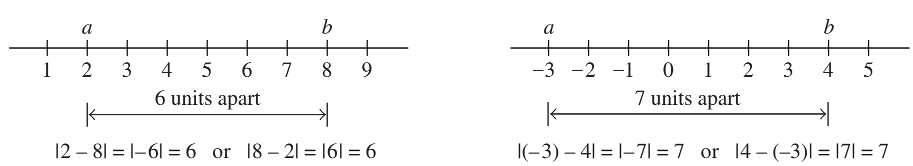


FIGURE 3 Distances in $\mathbb{R}$.

DEFINITION

For $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$, the **distance between u and v**, written as $\text{dist}(\mathbf{u}, \mathbf{v})$, is the length of the vector $\mathbf{u} - \mathbf{v}$. That is,

$$\text{dist}(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$$

In $\mathbb{R}^2$ and $\mathbb{R}^3$, this definition of distance coincides with the usual formulas for the Euclidean distance between two points, as the next two examples show.

EXAMPLE 4 Compute the distance between the vectors $\mathbf{u} = (7, 1)$ and $\mathbf{v} = (3, 2)$.

SOLUTION Calculate

$$\mathbf{u} - \mathbf{v} = \begin{bmatrix} 7 \\ 1 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$$

$$\|\mathbf{u} - \mathbf{v}\| = \sqrt{4^2 + (-1)^2} = \sqrt{17}$$

The vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{u} - \mathbf{v}$ are shown in Fig. 4. When the vector $\mathbf{u} - \mathbf{v}$ is added to $\mathbf{v}$, the result is $\mathbf{u}$. Notice that the parallelogram in Fig. 4 shows that the distance from $\mathbf{u}$ to $\mathbf{v}$ is the same as the distance from $\mathbf{u} - \mathbf{v}$ to $\mathbf{0}$. ■

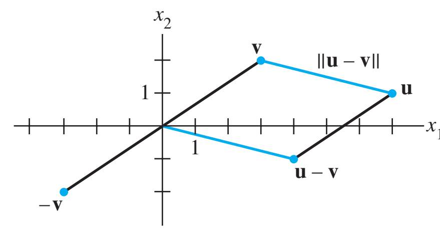


FIGURE 4 The distance between $\mathbf{u}$ and $\mathbf{v}$ is the length of $\mathbf{u} - \mathbf{v}$.

EXAMPLE 5 If $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (v_1, v_2, v_3)$, then

$$\begin{aligned} \text{dist}(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| &= \sqrt{(\mathbf{u} - \mathbf{v}) \cdot (\mathbf{u} - \mathbf{v})} \\ &= \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + (u_3 - v_3)^2} \end{aligned}$$ ■

Orthogonal Vectors

The rest of this chapter depends on the fact that the concept of perpendicular lines in ordinary Euclidean geometry has an analogue in $\mathbb{R}^n$.

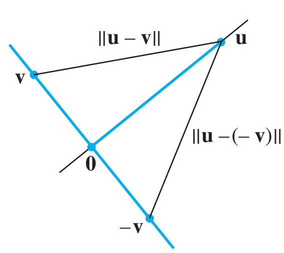


FIGURE 5

Consider $\mathbb{R}^2$ or $\mathbb{R}^3$ and two lines through the origin determined by vectors $\mathbf{u}$ and $\mathbf{v}$. The two lines shown in Fig. 5 are geometrically perpendicular if and only if the distance from $\mathbf{u}$ to $\mathbf{v}$ is the same as the distance from $\mathbf{u}$ to $-\mathbf{v}$. This is the same as requiring the squares of the distances to be the same. Now

$$\begin{aligned} [\,\text{dist}(\mathbf{u}, -\mathbf{v})\,]^2 &= \|\mathbf{u} - (-\mathbf{v})\|^2 = \|\mathbf{u} + \mathbf{v}\|^2 \\ &= (\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v}) \\ &= \mathbf{u} \cdot (\mathbf{u} + \mathbf{v}) + \mathbf{v} \cdot (\mathbf{u} + \mathbf{v}) && \text{Theorem 1(b)} \\ &= \mathbf{u} \cdot \mathbf{u} + \mathbf{u} \cdot \mathbf{v} + \mathbf{v} \cdot \mathbf{u} + \mathbf{v} \cdot \mathbf{v} && \text{Theorem 1(a), (b)} \\ &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\mathbf{u} \cdot \mathbf{v} && \text{Theorem 1(a)} \end{aligned} \tag{1}$$

The same calculations with $\mathbf{v}$ and $-\mathbf{v}$ interchanged show that

$$\begin{aligned}[\text{dist}(\mathbf{u}, \mathbf{v})]^2 &= \|\mathbf{u}\|^2 + \|-\mathbf{v}\|^2 + 2\mathbf{u}\cdot(-\mathbf{v}) \\ &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\mathbf{u}\cdot\mathbf{v}\end{aligned}$$

The two squared distances are equal if and only if $2\mathbf{u}\cdot\mathbf{v} = -2\mathbf{u}\cdot\mathbf{v}$, which happens if and only if $\mathbf{u}\cdot\mathbf{v} = 0$.

This calculation shows that when vectors $\mathbf{u}$ and $\mathbf{v}$ are identified with geometric points, the corresponding lines through the points and the origin are perpendicular if and only if $\mathbf{u}\cdot\mathbf{v} = 0$. The following definition generalizes to $\mathbb{R}^n$ this notion of perpendicularity (or *orthogonality*, as it is commonly called in linear algebra).

DEFINITION

Two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ are **orthogonal** (to each other) if $\mathbf{u}\cdot\mathbf{v} = 0$.

Observe that the zero vector is orthogonal to every vector in $\mathbb{R}^n$ because $\mathbf{0}^T\mathbf{v} = 0$ for all $\mathbf{v}$.

The next theorem provides a useful fact about orthogonal vectors. The proof follows immediately from the calculation in (1) above and the definition of orthogonality. The right triangle shown in Fig. 6 provides a visualization of the lengths that appear in the theorem.

THEOREM 2

The Pythagorean Theorem

Two vectors $\mathbf{u}$ and $\mathbf{v}$ are orthogonal if and only if $\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$.

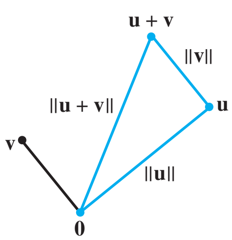


FIGURE 6

Orthogonal Complements

To provide practice using inner products, we introduce a concept here that will be of use in Section 6.3 and elsewhere in the chapter. If a vector $\mathbf{z}$ is orthogonal to every vector in a subspace W of $\mathbb{R}^n$, then $\mathbf{z}$ is said to be **orthogonal to** W. The set of all vectors $\mathbf{z}$ that are orthogonal to W is called the **orthogonal complement** of W and is denoted by $W^\perp$ (and read as "W perpendicular" or simply "W perp").

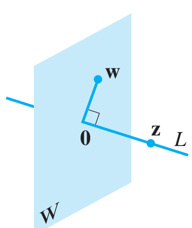


FIGURE 7
A plane and line through **0** as orthogonal complements.

EXAMPLE 6 Let W be a plane through the origin in $\mathbb{R}^3$, and let L be the line through the origin and perpendicular to W. If $\mathbf{z}$ and $\mathbf{w}$ are nonzero, $\mathbf{z}$ is on L, and $\mathbf{w}$ is in W, then the line segment from $\mathbf{0}$ to $\mathbf{z}$ is perpendicular to the line segment from $\mathbf{0}$ to $\mathbf{w}$; that is, $\mathbf{z}\cdot\mathbf{w} = 0$. See Fig. 7. So each vector on L is orthogonal to every $\mathbf{w}$ in W. In fact, L consists of *all* vectors that are orthogonal to the $\mathbf{w}$'s in W, and W consists of all vectors orthogonal to the $\mathbf{z}$'s in L. That is,

$$L = W^\perp \quad \text{and} \quad W = L^\perp$$ ■

The following two facts about $W^\perp$, with W a subspace of $\mathbb{R}^n$, are needed later in the chapter. Proofs are suggested in Exercises 29 and 30. Exercises 27–31 provide excellent practice using properties of the inner product.

1. A vector $\mathbf{x}$ is in $W^\perp$ if and only if $\mathbf{x}$ is orthogonal to every vector in a set that spans W.
2. $W^\perp$ is a subspace of $\mathbb{R}^n$.

The next theorem and Exercise 31 verify the claims made in Section 4.6 concerning the subspaces shown in Fig. 8. (Also see Exercise 28 in Section 4.6.)

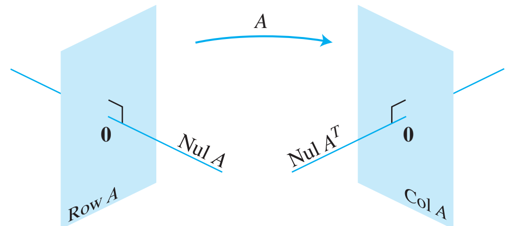


FIGURE 8 The fundamental subspaces determined by an $m \times n$ matrix A.

THEOREM 3

Let A be an $m \times n$ matrix. The orthogonal complement of the row space of A is the null space of A, and the orthogonal complement of the column space of A is the null space of A^T:

$$(\text{Row } A)^{\perp} = \text{Nul } A \quad \text{and} \quad (\text{Col } A)^{\perp} = \text{Nul } A^T$$

PROOF The row–column rule for computing $A\mathbf{x}$ shows that if $\mathbf{x}$ is in Nul A, then $\mathbf{x}$ is orthogonal to each row of A (with the rows treated as vectors in $\mathbb{R}^n$). Since the rows of A span the row space, $\mathbf{x}$ is orthogonal to Row A. Conversely, if $\mathbf{x}$ is orthogonal to Row A, then $\mathbf{x}$ is certainly orthogonal to each row of A, and hence $A\mathbf{x} = \mathbf{0}$. This proves the first statement of the theorem. Since this statement is true for any matrix, it is true for A^T. That is, the orthogonal complement of the row space of A^T is the null space of A^T. This proves the second statement, because Row A^T = Col A. ■

Angles in $\mathbb{R}^2$ and $\mathbb{R}^3$ (Optional)

If $\mathbf{u}$ and $\mathbf{v}$ are nonzero vectors in either $\mathbb{R}^2$ or $\mathbb{R}^3$, then there is a nice connection between their inner product and the angle ϑ between the two line segments from the origin to the points identified with $\mathbf{u}$ and $\mathbf{v}$. The formula is

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \, \|\mathbf{v}\| \cos \vartheta \tag{2}$$

To verify this formula for vectors in $\mathbb{R}^2$, consider the triangle shown in Fig. 9, with sides of lengths $\|\mathbf{u}\|$, $\|\mathbf{v}\|$, and $\|\mathbf{u} - \mathbf{v}\|$. By the law of cosines,

$$\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\| \, \|\mathbf{v}\| \cos \vartheta$$

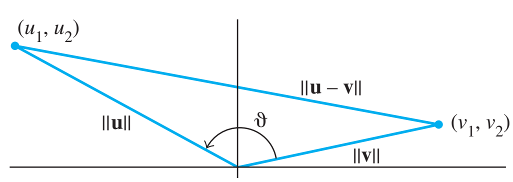


FIGURE 9 The angle between two vectors.

which can be rearranged to produce

$$\begin{aligned}\|\mathbf{u}\|\,\|\mathbf{v}\|\cos\vartheta &= \frac{1}{2}\left[\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - \|\mathbf{u}-\mathbf{v}\|^2\right] \\ &= \frac{1}{2}\left[u_1^2 + u_2^2 + v_1^2 + v_2^2 - (u_1-v_1)^2 - (u_2-v_2)^2\right] \\ &= u_1v_1 + u_2v_2 \\ &= \mathbf{u}\cdot\mathbf{v}\end{aligned}$$

The verification for $\mathbb{R}^3$ is similar. When $n > 3$, formula (2) may be used to *define* the angle between two vectors in $\mathbb{R}^n$. In statistics, for instance, the value of $\cos\vartheta$ defined by (2) for suitable vectors **u** and **v** is what statisticians call a *correlation coefficient*.

PRACTICE PROBLEMS

Let $\mathbf{a} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$, $\mathbf{c} = \begin{bmatrix} 4/3 \\ -1 \\ 2/3 \end{bmatrix}$, and $\mathbf{d} = \begin{bmatrix} 5 \\ 6 \\ -1 \end{bmatrix}$.

1. Compute $\dfrac{\mathbf{a}\cdot\mathbf{b}}{\mathbf{a}\cdot\mathbf{a}}$ and $\left(\dfrac{\mathbf{a}\cdot\mathbf{b}}{\mathbf{a}\cdot\mathbf{a}}\right)\mathbf{a}$.
2. Find a unit vector **u** in the direction of **c**.
3. Show that **d** is orthogonal to **c**.
4. Use the results of Practice Problems 2 and 3 to explain why **d** must be orthogonal to the unit vector **u**.

6.1 EXERCISES

Compute the quantities in Exercises 1–8 using the vectors

$$\mathbf{u} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} 3 \\ -1 \\ -5 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} 6 \\ -2 \\ 3 \end{bmatrix}$$

1. $\mathbf{u}\cdot\mathbf{u}$, $\mathbf{v}\cdot\mathbf{u}$, and $\dfrac{\mathbf{v}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}$

2. $\mathbf{w}\cdot\mathbf{w}$, $\mathbf{x}\cdot\mathbf{w}$, and $\dfrac{\mathbf{x}\cdot\mathbf{w}}{\mathbf{w}\cdot\mathbf{w}}$

3. $\dfrac{1}{\mathbf{w}\cdot\mathbf{w}}\mathbf{w}$

4. $\dfrac{1}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}$

5. $\left(\dfrac{\mathbf{u}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\right)\mathbf{v}$

6. $\left(\dfrac{\mathbf{x}\cdot\mathbf{w}}{\mathbf{x}\cdot\mathbf{x}}\right)\mathbf{x}$

7. $\|\mathbf{w}\|$

8. $\|\mathbf{x}\|$

In Exercises 9–12, find a unit vector in the direction of the given vector.

9. $\begin{bmatrix} -30 \\ 40 \end{bmatrix}$

10. $\begin{bmatrix} -6 \\ 4 \\ -3 \end{bmatrix}$

11. $\begin{bmatrix} 7/4 \\ 1/2 \\ 1 \end{bmatrix}$

12. $\begin{bmatrix} 8/3 \\ 2 \end{bmatrix}$

13. Find the distance between $\mathbf{x} = \begin{bmatrix} 10 \\ -3 \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} -1 \\ -5 \end{bmatrix}$.

14. Find the distance between $\mathbf{u} = \begin{bmatrix} 0 \\ -5 \\ 2 \end{bmatrix}$ and $\mathbf{z} = \begin{bmatrix} -4 \\ -1 \\ 8 \end{bmatrix}$.

Determine which pairs of vectors in Exercises 15–18 are orthogonal.

15. $\mathbf{a} = \begin{bmatrix} 8 \\ -5 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} -2 \\ -3 \end{bmatrix}$

16. $\mathbf{u} = \begin{bmatrix} 12 \\ 3 \\ -5 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 2 \\ -3 \\ 3 \end{bmatrix}$

17. $\mathbf{u} = \begin{bmatrix} 3 \\ 2 \\ -5 \\ 0 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} -4 \\ 1 \\ -2 \\ 6 \end{bmatrix}$

18. $\mathbf{y} = \begin{bmatrix} -3 \\ 7 \\ 4 \\ 0 \end{bmatrix}$, $\mathbf{z} = \begin{bmatrix} 1 \\ -8 \\ 15 \\ -7 \end{bmatrix}$

In Exercises 19 and 20, all vectors are in $\mathbb{R}^n$. Mark each statement True or False. Justify each answer.

19. a. $\mathbf{v}\cdot\mathbf{v} = \|\mathbf{v}\|^2$.
    b. For any scalar c, $\mathbf{u}\cdot(c\mathbf{v}) = c(\mathbf{u}\cdot\mathbf{v})$.
    c. If the distance from **u** to **v** equals the distance from **u** to $-\mathbf{v}$, then **u** and **v** are orthogonal.
    d. For a square matrix A, vectors in Col A are orthogonal to vectors in Nul A.
    e. If vectors $\mathbf{v}_1, \dots, \mathbf{v}_p$ span a subspace W and if **x** is orthogonal to each $\mathbf{v}_j$ for $j = 1, \dots, p$, then **x** is in $W^{\perp}$.

20. a. $\mathbf{u}\cdot\mathbf{v} - \mathbf{v}\cdot\mathbf{u} = 0$.

b. For any scalar c, $\|c\mathbf{v}\| = c\|\mathbf{v}\|$.

c. If $\mathbf{x}$ is orthogonal to every vector in a subspace W, then $\mathbf{x}$ is in $W^{\perp}$.

d. If $\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 = \|\mathbf{u}+\mathbf{v}\|^2$, then $\mathbf{u}$ and $\mathbf{v}$ are orthogonal.

e. For an $m \times n$ matrix A, vectors in the null space of A are orthogonal to vectors in the row space of A.

21. Use the transpose definition of the inner product to verify parts (b) and (c) of Theorem 1. Mention the appropriate facts from Chapter 2.

22. Let $\mathbf{u} = (u_1, u_2, u_3)$. Explain why $\mathbf{u}\cdot\mathbf{u} \geq 0$. When is $\mathbf{u}\cdot\mathbf{u} = 0$?

23. Let $\mathbf{u} = \begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} -7 \\ -4 \\ 6 \end{bmatrix}$. Compute and compare $\mathbf{u}\cdot\mathbf{v}$, $\|\mathbf{u}\|^2$, $\|\mathbf{v}\|^2$, and $\|\mathbf{u}+\mathbf{v}\|^2$. Do not use the Pythagorean Theorem.

24. Verify the *parallelogram law* for vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$:

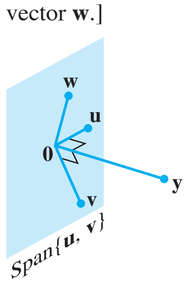
$$\|\mathbf{u}+\mathbf{v}\|^2 + \|\mathbf{u}-\mathbf{v}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2$$

25. Let $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$. Describe the set H of vectors $\begin{bmatrix} x \\ y \end{bmatrix}$ that are orthogonal to $\mathbf{v}$. [*Hint:* Consider $\mathbf{v} = \mathbf{0}$ and $\mathbf{v} \neq \mathbf{0}$.]

26. Let $\mathbf{u} = \begin{bmatrix} 5 \\ -6 \\ 7 \end{bmatrix}$, and let W be the set of all $\mathbf{x}$ in $\mathbb{R}^3$ such that $\mathbf{u}\cdot\mathbf{x} = 0$. What theorem in Chapter 4 can be used to show that W is a subspace of $\mathbb{R}^3$? Describe W in geometric language.

27. Suppose a vector $\mathbf{y}$ is orthogonal to vectors $\mathbf{u}$ and $\mathbf{v}$. Show that $\mathbf{y}$ is orthogonal to the vector $\mathbf{u}+\mathbf{v}$.

28. Suppose $\mathbf{y}$ is orthogonal to $\mathbf{u}$ and $\mathbf{v}$. Show that $\mathbf{y}$ is orthogonal to every $\mathbf{w}$ in Span $\{\mathbf{u}, \mathbf{v}\}$. [*Hint:* An arbitrary $\mathbf{w}$ in Span $\{\mathbf{u}, \mathbf{v}\}$ has the form $\mathbf{w} = c_1\mathbf{u} + c_2\mathbf{v}$. Show that $\mathbf{y}$ is orthogonal to such a vector $\mathbf{w}$.]



29. Let $W = \text{Span}\,\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$. Show that if $\mathbf{x}$ is orthogonal to each $\mathbf{v}_j$, for $1 \leq j \leq p$, then $\mathbf{x}$ is orthogonal to every vector in W.

30. Let W be a subspace of $\mathbb{R}^n$, and let $W^{\perp}$ be the set of all vectors orthogonal to W. Show that $W^{\perp}$ is a subspace of $\mathbb{R}^n$ using the following steps.

a. Take $\mathbf{z}$ in $W^{\perp}$, and let $\mathbf{u}$ represent any element of W. Then $\mathbf{z}\cdot\mathbf{u} = 0$. Take any scalar c and show that $c\mathbf{z}$ is orthogonal to $\mathbf{u}$. (Since $\mathbf{u}$ was an arbitrary element of W, this will show that $c\mathbf{z}$ is in $W^{\perp}$.)

b. Take $\mathbf{z}_1$ and $\mathbf{z}_2$ in $W^{\perp}$, and let $\mathbf{u}$ be any element of W. Show that $\mathbf{z}_1 + \mathbf{z}_2$ is orthogonal to $\mathbf{u}$. What can you conclude about $\mathbf{z}_1 + \mathbf{z}_2$? Why?

c. Finish the proof that $W^{\perp}$ is a subspace of $\mathbb{R}^n$.

31. Show that if $\mathbf{x}$ is in both W and $W^{\perp}$, then $\mathbf{x} = \mathbf{0}$.

32. **[M]** Construct a pair $\mathbf{u}, \mathbf{v}$ of random vectors in $\mathbb{R}^4$, and let

$$A = \begin{bmatrix} .5 & .5 & .5 & .5 \\ .5 & .5 & -.5 & -.5 \\ .5 & -.5 & .5 & -.5 \\ .5 & -.5 & -.5 & .5 \end{bmatrix}$$

a. Denote the columns of A by $\mathbf{a}_1, \ldots, \mathbf{a}_4$. Compute the length of each column, and compute $\mathbf{a}_1\cdot\mathbf{a}_2$, $\mathbf{a}_1\cdot\mathbf{a}_3$, $\mathbf{a}_1\cdot\mathbf{a}_4$, $\mathbf{a}_2\cdot\mathbf{a}_3$, $\mathbf{a}_2\cdot\mathbf{a}_4$, and $\mathbf{a}_3\cdot\mathbf{a}_4$.

b. Compute and compare the lengths of $\mathbf{u}$, $A\mathbf{u}$, $\mathbf{v}$, and $A\mathbf{v}$.

c. Use equation (2) in this section to compute the cosine of the angle between $\mathbf{u}$ and $\mathbf{v}$. Compare this with the cosine of the angle between $A\mathbf{u}$ and $A\mathbf{v}$.

d. Repeat parts (b) and (c) for two other pairs of random vectors. What do you conjecture about the effect of A on vectors?

33. **[M]** Generate random vectors $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{v}$ in $\mathbb{R}^4$ with integer entries (and $\mathbf{v} \neq \mathbf{0}$), and compute the quantities

$$\left(\frac{\mathbf{x}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\right)\mathbf{v},\ \left(\frac{\mathbf{y}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\right)\mathbf{v},\ \frac{(\mathbf{x}+\mathbf{y})\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v},\ \frac{(10\mathbf{x})\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v}$$

Repeat the computations with new random vectors $\mathbf{x}$ and $\mathbf{y}$. What do you conjecture about the mapping $\mathbf{x} \mapsto T(\mathbf{x}) = \left(\frac{\mathbf{x}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\right)\mathbf{v}$ (for $\mathbf{v} \neq \mathbf{0}$)? Verify your conjecture algebraically.

34. **[M]** Let $A = \begin{bmatrix} -6 & 3 & -27 & -33 & -13 \\ 6 & -5 & 25 & 28 & 14 \\ 8 & -6 & 34 & 38 & 18 \\ 12 & -10 & 50 & 41 & 23 \\ 14 & -21 & 49 & 29 & 33 \end{bmatrix}$. Construct a matrix N whose columns form a basis for Nul A, and construct a matrix R whose *rows* form a basis for Row A (see Section 4.6 for details). Perform a matrix computation with N and R that illustrates a fact from Theorem 3.

SOLUTIONS TO PRACTICE PROBLEMS

1. $\mathbf{a}\cdot\mathbf{b} = 7$, $\mathbf{a}\cdot\mathbf{a} = 5$. Hence $\dfrac{\mathbf{a}\cdot\mathbf{b}}{\mathbf{a}\cdot\mathbf{a}} = \dfrac{7}{5}$, and $\left(\dfrac{\mathbf{a}\cdot\mathbf{b}}{\mathbf{a}\cdot\mathbf{a}}\right)\mathbf{a} = \dfrac{7}{5}\mathbf{a} = \begin{bmatrix} -14/5 \\ 7/5 \end{bmatrix}$.

2. Scale $\mathbf{c}$, multiplying by 3 to get $\mathbf{y} = \begin{bmatrix} 4 \\ -3 \\ 2 \end{bmatrix}$. Compute $\|\mathbf{y}\|^2 = 29$ and $\|\mathbf{y}\| = \sqrt{29}$.

   The unit vector in the direction of both $\mathbf{c}$ and $\mathbf{y}$ is $\mathbf{u} = \dfrac{1}{\|\mathbf{y}\|}\mathbf{y} = \begin{bmatrix} 4/\sqrt{29} \\ -3/\sqrt{29} \\ 2/\sqrt{29} \end{bmatrix}$.

3. $\mathbf{d}$ is orthogonal to $\mathbf{c}$, because

$$\mathbf{d}\cdot\mathbf{c} = \begin{bmatrix} 5 \\ 6 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} 4/3 \\ -1 \\ 2/3 \end{bmatrix} = \frac{20}{3} - 6 - \frac{2}{3} = 0$$

4. $\mathbf{d}$ is orthogonal to $\mathbf{u}$ because $\mathbf{u}$ has the form $k\mathbf{c}$ for some k, and

$$\mathbf{d}\cdot\mathbf{u} = \mathbf{d}\cdot(k\mathbf{c}) = k(\mathbf{d}\cdot\mathbf{c}) = k(0) = 0$$

6.2 ORTHOGONAL SETS

A set of vectors $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ in $\mathbb{R}^n$ is said to be an **orthogonal set** if each pair of distinct vectors from the set is orthogonal, that is, if $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ whenever $i \neq j$.

EXAMPLE 1 Show that $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal set, where

$$\mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix}$$

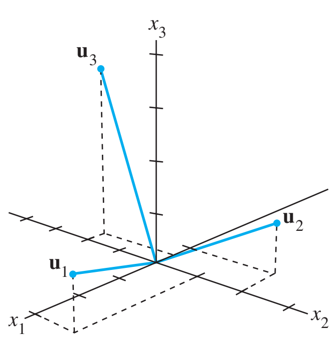


FIGURE 1

SOLUTION Consider the three possible pairs of distinct vectors, namely, $\{\mathbf{u}_1, \mathbf{u}_2\}$, $\{\mathbf{u}_1, \mathbf{u}_3\}$, and $\{\mathbf{u}_2, \mathbf{u}_3\}$.

$$\begin{aligned}
\mathbf{u}_1\cdot\mathbf{u}_2 &= 3(-1) + 1(2) + 1(1) = 0 \\
\mathbf{u}_1\cdot\mathbf{u}_3 &= 3\left(-\tfrac{1}{2}\right) + 1(-2) + 1\left(\tfrac{7}{2}\right) = 0 \\
\mathbf{u}_2\cdot\mathbf{u}_3 &= -1\left(-\tfrac{1}{2}\right) + 2(-2) + 1\left(\tfrac{7}{2}\right) = 0
\end{aligned}$$

Each pair of distinct vectors is orthogonal, and so $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal set. See Fig. 1; the three line segments there are mutually perpendicular. ■

THEOREM 4

If $S = \{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is an orthogonal set of nonzero vectors in $\mathbb{R}^n$, then S is linearly independent and hence is a basis for the subspace spanned by S.

PROOF If $\mathbf{0} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$ for some scalars $c_1, \dots, c_p$, then

$$\begin{aligned}
0 = \mathbf{0}\cdot\mathbf{u}_1 &= (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p)\cdot\mathbf{u}_1 \\
&= (c_1\mathbf{u}_1)\cdot\mathbf{u}_1 + (c_2\mathbf{u}_2)\cdot\mathbf{u}_1 + \cdots + (c_p\mathbf{u}_p)\cdot\mathbf{u}_1 \\
&= c_1(\mathbf{u}_1\cdot\mathbf{u}_1) + c_2(\mathbf{u}_2\cdot\mathbf{u}_1) + \cdots + c_p(\mathbf{u}_p\cdot\mathbf{u}_1) \\
&= c_1(\mathbf{u}_1\cdot\mathbf{u}_1)
\end{aligned}$$

because $\mathbf{u}_1$ is orthogonal to $\mathbf{u}_2, \dots, \mathbf{u}_p$. Since $\mathbf{u}_1$ is nonzero, $\mathbf{u}_1\cdot\mathbf{u}_1$ is not zero and so $c_1 = 0$. Similarly, $c_2, \dots, c_p$ must be zero. Thus S is linearly independent. ■

DEFINITION

An **orthogonal basis** for a subspace W of $\mathbb{R}^n$ is a basis for W that is also an orthogonal set.

The next theorem suggests why an orthogonal basis is much nicer than other bases. The weights in a linear combination can be computed easily.

THEOREM 5

Let $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ be an orthogonal basis for a subspace W of $\mathbb{R}^n$. For each $\mathbf{y}$ in W, the weights in the linear combination

$$\mathbf{y} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$$

are given by

$$c_j = \frac{\mathbf{y}\cdot\mathbf{u}_j}{\mathbf{u}_j\cdot\mathbf{u}_j} \qquad (j = 1, \dots, p)$$

PROOF As in the preceding proof, the orthogonality of $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ shows that

$$\mathbf{y}\cdot\mathbf{u}_1 = (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p)\cdot\mathbf{u}_1 = c_1(\mathbf{u}_1\cdot\mathbf{u}_1)$$

Since $\mathbf{u}_1\cdot\mathbf{u}_1$ is not zero, the equation above can be solved for c_1. To find c_j for $j = 2, \dots, p$, compute $\mathbf{y}\cdot\mathbf{u}_j$ and solve for c_j. ■

EXAMPLE 2 The set $S = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ in Example 1 is an orthogonal basis for $\mathbb{R}^3$. Express the vector $\mathbf{y} = \begin{bmatrix} 6 \\ 1 \\ -8 \end{bmatrix}$ as a linear combination of the vectors in S.

SOLUTION Compute

$$\begin{aligned} \mathbf{y}\cdot\mathbf{u}_1 &= 11, & \mathbf{y}\cdot\mathbf{u}_2 &= -12, & \mathbf{y}\cdot\mathbf{u}_3 &= -33 \\ \mathbf{u}_1\cdot\mathbf{u}_1 &= 11, & \mathbf{u}_2\cdot\mathbf{u}_2 &= 6, & \mathbf{u}_3\cdot\mathbf{u}_3 &= 33/2 \end{aligned}$$

By Theorem 5,

$$\begin{aligned} \mathbf{y} &= \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2 + \frac{\mathbf{y}\cdot\mathbf{u}_3}{\mathbf{u}_3\cdot\mathbf{u}_3}\mathbf{u}_3 \\ &= \frac{11}{11}\mathbf{u}_1 + \frac{-12}{6}\mathbf{u}_2 + \frac{-33}{33/2}\mathbf{u}_3 \\ &= \mathbf{u}_1 - 2\mathbf{u}_2 - 2\mathbf{u}_3 \end{aligned}$$

■

Notice how easy it is to compute the weights needed to build $\mathbf{y}$ from an orthogonal basis. If the basis were not orthogonal, it would be necessary to solve a system of linear equations in order to find the weights, as in Chapter 1.

We turn next to a construction that will become a key step in many calculations involving orthogonality, and it will lead to a geometric interpretation of Theorem 5.

An Orthogonal Projection

Given a nonzero vector $\mathbf{u}$ in $\mathbb{R}^n$, consider the problem of decomposing a vector $\mathbf{y}$ in $\mathbb{R}^n$ into the sum of two vectors, one a multiple of $\mathbf{u}$ and the other orthogonal to $\mathbf{u}$. We wish to write

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} \tag{1}$$

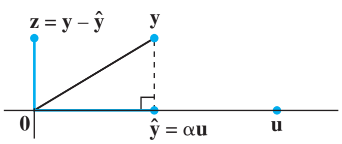


FIGURE 2 Finding α to make $\mathbf{y} - \hat{\mathbf{y}}$ orthogonal to $\mathbf{u}$.

where $\hat{\mathbf{y}} = \alpha\mathbf{u}$ for some scalar α and $\mathbf{z}$ is some vector orthogonal to $\mathbf{u}$. See Fig. 2. Given any scalar α, let $\mathbf{z} = \mathbf{y} - \alpha\mathbf{u}$, so that (1) is satisfied. Then $\mathbf{y} - \hat{\mathbf{y}}$ is orthogonal to $\mathbf{u}$ if and only if

$$0 = (\mathbf{y} - \alpha\mathbf{u})\cdot\mathbf{u} = \mathbf{y}\cdot\mathbf{u} - (\alpha\mathbf{u})\cdot\mathbf{u} = \mathbf{y}\cdot\mathbf{u} - \alpha(\mathbf{u}\cdot\mathbf{u})$$

That is, (1) is satisfied with $\mathbf{z}$ orthogonal to $\mathbf{u}$ if and only if $\alpha = \dfrac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}$ and $\hat{\mathbf{y}} = \dfrac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}$. The vector $\hat{\mathbf{y}}$ is called the **orthogonal projection of y onto u**, and the vector $\mathbf{z}$ is called the **component of y orthogonal to u**.

If c is any nonzero scalar and if $\mathbf{u}$ is replaced by $c\mathbf{u}$ in the definition of $\hat{\mathbf{y}}$, then the orthogonal projection of $\mathbf{y}$ onto $c\mathbf{u}$ is exactly the same as the orthogonal projection of $\mathbf{y}$ onto $\mathbf{u}$ (Exercise 31). Hence this projection is determined by the *subspace* L spanned by $\mathbf{u}$ (the line through $\mathbf{u}$ and $\mathbf{0}$). Sometimes $\hat{\mathbf{y}}$ is denoted by $\text{proj}_L\,\mathbf{y}$ and is called the **orthogonal projection of y onto** L. That is,

$$\boxed{\hat{\mathbf{y}} = \text{proj}_L\,\mathbf{y} = \frac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}} \tag{2}$$

EXAMPLE 3 Let $\mathbf{y} = \begin{bmatrix} 7 \\ 6 \end{bmatrix}$ and $\mathbf{u} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$. Find the orthogonal projection of $\mathbf{y}$ onto $\mathbf{u}$. Then write $\mathbf{y}$ as the sum of two orthogonal vectors, one in Span $\{\mathbf{u}\}$ and one orthogonal to $\mathbf{u}$.

SOLUTION Compute

$$\mathbf{y}\cdot\mathbf{u} = \begin{bmatrix} 7 \\ 6 \end{bmatrix}\cdot\begin{bmatrix} 4 \\ 2 \end{bmatrix} = 40$$

$$\mathbf{u}\cdot\mathbf{u} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}\cdot\begin{bmatrix} 4 \\ 2 \end{bmatrix} = 20$$

The orthogonal projection of $\mathbf{y}$ onto $\mathbf{u}$ is

$$\hat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u} = \frac{40}{20}\mathbf{u} = 2\begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

and the component of $\mathbf{y}$ orthogonal to $\mathbf{u}$ is

$$\mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} 7 \\ 6 \end{bmatrix} - \begin{bmatrix} 8 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

The sum of these two vectors is $\mathbf{y}$. That is,

$$\underset{\substack{\uparrow \\ \mathbf{y}}}{\begin{bmatrix} 7 \\ 6 \end{bmatrix}} = \underset{\substack{\uparrow \\ \hat{\mathbf{y}}}}{\begin{bmatrix} 8 \\ 4 \end{bmatrix}} + \underset{\substack{\uparrow \\ (\mathbf{y} - \hat{\mathbf{y}})}}{\begin{bmatrix} -1 \\ 2 \end{bmatrix}}$$

This decomposition of $\mathbf{y}$ is illustrated in Fig. 3. *Note:* If the calculations above are correct, then $\{\hat{\mathbf{y}}, \mathbf{y} - \hat{\mathbf{y}}\}$ will be an orthogonal set. As a check, compute

$$\hat{\mathbf{y}}\cdot(\mathbf{y} - \hat{\mathbf{y}}) = \begin{bmatrix} 8 \\ 4 \end{bmatrix}\cdot\begin{bmatrix} -1 \\ 2 \end{bmatrix} = -8 + 8 = 0$$ ■

Since the line segment in Fig. 3 between $\mathbf{y}$ and $\hat{\mathbf{y}}$ is perpendicular to L, by construction of $\hat{\mathbf{y}}$, the point identified with $\hat{\mathbf{y}}$ is the closest point of L to $\mathbf{y}$. (This can be proved from geometry. We will assume this for $\mathbb{R}^2$ now and prove it for $\mathbb{R}^n$ in Section 6.3.)

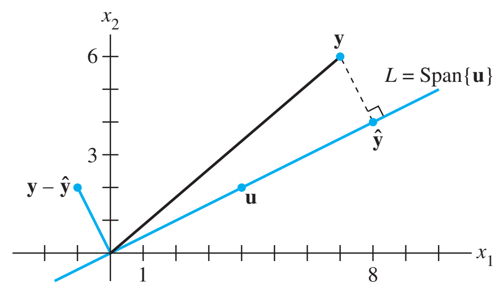


FIGURE 3 The orthogonal projection of $\mathbf{y}$ onto a line L through the origin.

EXAMPLE 4 Find the distance in Fig. 3 from $\mathbf{y}$ to L.

SOLUTION The distance from $\mathbf{y}$ to L is the length of the perpendicular line segment from $\mathbf{y}$ to the orthogonal projection $\hat{\mathbf{y}}$. This length equals the length of $\mathbf{y} - \hat{\mathbf{y}}$. Thus the distance is

$$\|\mathbf{y} - \hat{\mathbf{y}}\| = \sqrt{(-1)^2 + 2^2} = \sqrt{5}$$

A Geometric Interpretation of Theorem 5

The formula for the orthogonal projection $\hat{\mathbf{y}}$ in (2) has the same appearance as each of the terms in Theorem 5. Thus Theorem 5 decomposes a vector $\mathbf{y}$ into a sum of orthogonal projections onto one-dimensional subspaces.

It is easy to visualize the case in which $W = \mathbb{R}^2 = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$, with $\mathbf{u}_1$ and $\mathbf{u}_2$ orthogonal. Any $\mathbf{y}$ in $\mathbb{R}^2$ can be written in the form

$$\mathbf{y} = \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 \tag{3}$$

The first term in (3) is the projection of $\mathbf{y}$ onto the subspace spanned by $\mathbf{u}_1$ (the line through $\mathbf{u}_1$ and the origin), and the second term is the projection of $\mathbf{y}$ onto the subspace spanned by $\mathbf{u}_2$. Thus (3) expresses $\mathbf{y}$ as the sum of its projections onto the (orthogonal) axes determined by $\mathbf{u}_1$ and $\mathbf{u}_2$. See Fig. 4.

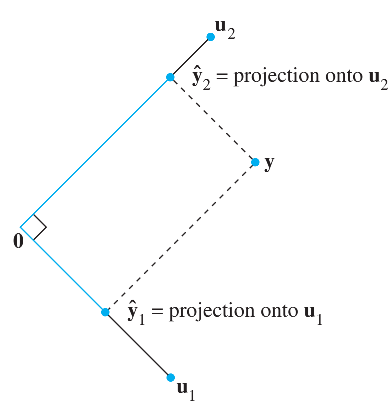


FIGURE 4 A vector decomposed into the sum of two projections.

Theorem 5 decomposes each $\mathbf{y}$ in $\text{Span}\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ into the sum of p projections onto one-dimensional subspaces that are mutually orthogonal.

Decomposing a Force into Component Forces

The decomposition in Fig. 4 can occur in physics when some sort of force is applied to an object. Choosing an appropriate coordinate system allows the force to be represented by a vector $\mathbf{y}$ in $\mathbb{R}^2$ or $\mathbb{R}^3$. Often the problem involves some particular direction of interest, which is represented by another vector $\mathbf{u}$. For instance, if the object is moving in a straight line when the force is applied, the vector $\mathbf{u}$ might point in the direction of movement, as in Fig. 5. A key step in the problem is to decompose the force into a component in the direction of $\mathbf{u}$ and a component orthogonal to $\mathbf{u}$. The calculations would be analogous to those made in Example 3 above.

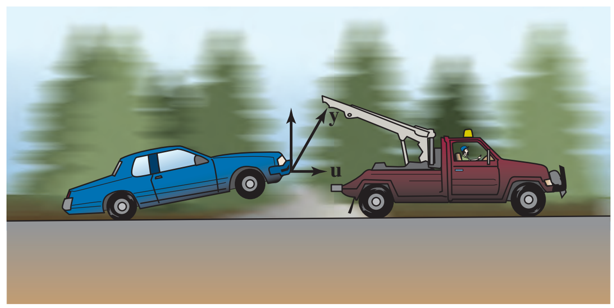


FIGURE 5

Orthonormal Sets

A set $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is an **orthonormal set** if it is an orthogonal set of unit vectors. If W is the subspace spanned by such a set, then $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is an **orthonormal basis** for W, since the set is automatically linearly independent, by Theorem 4.

The simplest example of an orthonormal set is the standard basis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ for $\mathbb{R}^n$. Any nonempty subset of $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ is orthonormal, too. Here is a more complicated example.

EXAMPLE 5 Show that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthonormal basis of $\mathbb{R}^3$, where

$$\mathbf{v}_1 = \begin{bmatrix} 3/\sqrt{11} \\ 1/\sqrt{11} \\ 1/\sqrt{11} \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -1/\sqrt{66} \\ -4/\sqrt{66} \\ 7/\sqrt{66} \end{bmatrix}$$

SOLUTION Compute

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = -3/\sqrt{66} + 2/\sqrt{66} + 1/\sqrt{66} = 0$$
$$\mathbf{v}_1 \cdot \mathbf{v}_3 = -3/\sqrt{726} - 4/\sqrt{726} + 7/\sqrt{726} = 0$$
$$\mathbf{v}_2 \cdot \mathbf{v}_3 = 1/\sqrt{396} - 8/\sqrt{396} + 7/\sqrt{396} = 0$$

Thus $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal set. Also,

$$\mathbf{v}_1 \cdot \mathbf{v}_1 = 9/11 + 1/11 + 1/11 = 1$$
$$\mathbf{v}_2 \cdot \mathbf{v}_2 = 1/6 + 4/6 + 1/6 = 1$$
$$\mathbf{v}_3 \cdot \mathbf{v}_3 = 1/66 + 16/66 + 49/66 = 1$$

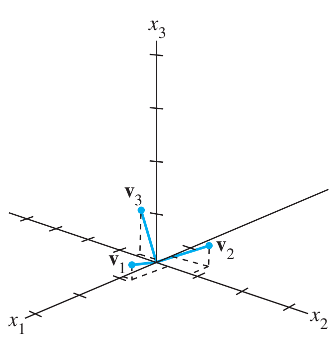


FIGURE 6

which shows that $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ are unit vectors. Thus $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthonormal set. Since the set is linearly independent, its three vectors form a basis for $\mathbb{R}^3$. See Fig. 6. ■

When the vectors in an orthogonal set of nonzero vectors are *normalized* to have unit length, the new vectors will still be orthogonal, and hence the new set will be an orthonormal set. See Exercise 32. It is easy to check that the vectors in Fig. 6 (Example 5) are simply the unit vectors in the directions of the vectors in Fig. 1 (Example 1).

Matrices whose columns form an orthonormal set are important in applications and in computer algorithms for matrix computations. Their main properties are given in Theorems 6 and 7.

THEOREM 6

An $m \times n$ matrix U has orthonormal columns if and only if $U^TU = I$.

PROOF To simplify notation, we suppose that U has only three columns, each a vector in $\mathbb{R}^m$. The proof of the general case is essentially the same. Let $U = [\,\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3\,]$ and compute

$$U^TU = \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \mathbf{u}_3^T \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1^T\mathbf{u}_1 & \mathbf{u}_1^T\mathbf{u}_2 & \mathbf{u}_1^T\mathbf{u}_3 \\ \mathbf{u}_2^T\mathbf{u}_1 & \mathbf{u}_2^T\mathbf{u}_2 & \mathbf{u}_2^T\mathbf{u}_3 \\ \mathbf{u}_3^T\mathbf{u}_1 & \mathbf{u}_3^T\mathbf{u}_2 & \mathbf{u}_3^T\mathbf{u}_3 \end{bmatrix} \tag{4}$$

The entries in the matrix at the right are inner products, using transpose notation. The columns of U are orthogonal if and only if

$$\mathbf{u}_1^T\mathbf{u}_2 = \mathbf{u}_2^T\mathbf{u}_1 = 0, \quad \mathbf{u}_1^T\mathbf{u}_3 = \mathbf{u}_3^T\mathbf{u}_1 = 0, \quad \mathbf{u}_2^T\mathbf{u}_3 = \mathbf{u}_3^T\mathbf{u}_2 = 0 \tag{5}$$

The columns of U all have unit length if and only if

$$\mathbf{u}_1^T\mathbf{u}_1 = 1, \quad \mathbf{u}_2^T\mathbf{u}_2 = 1, \quad \mathbf{u}_3^T\mathbf{u}_3 = 1 \tag{6}$$

The theorem follows immediately from (4)–(6). ■

THEOREM 7

Let U be an $m \times n$ matrix with orthonormal columns, and let $\mathbf{x}$ and $\mathbf{y}$ be in $\mathbb{R}^n$. Then

a. $\|U\mathbf{x}\| = \|\mathbf{x}\|$
b. $(U\mathbf{x})\cdot(U\mathbf{y}) = \mathbf{x}\cdot\mathbf{y}$
c. $(U\mathbf{x})\cdot(U\mathbf{y}) = 0$ if and only if $\mathbf{x}\cdot\mathbf{y} = 0$

Properties (a) and (c) say that the linear mapping $\mathbf{x} \mapsto U\mathbf{x}$ preserves lengths and orthogonality. These properties are crucial for many computer algorithms. See Exercise 25 for the proof of Theorem 7.

EXAMPLE 6 Let $U = \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix}$. Notice that U has orthonormal columns and

$$U^TU = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 2/3 & -2/3 & 1/3 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Verify that $\|U\mathbf{x}\| = \|\mathbf{x}\|$.

SOLUTION

$$U\mathbf{x} = \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix} \begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix}$$

$$\|U\mathbf{x}\| = \sqrt{9+1+1} = \sqrt{11}$$

$$\|\mathbf{x}\| = \sqrt{2+9} = \sqrt{11}$$

Theorems 6 and 7 are particularly useful when applied to *square* matrices. An **orthogonal matrix** is a square invertible matrix U such that $U^{-1} = U^T$. By Theorem 6, such a matrix has orthonormal columns.[1] It is easy to see that any *square* matrix with orthonormal columns is an orthogonal matrix. Surprisingly, such a matrix must have orthonormal *rows*, too. See Exercises 27 and 28. Orthogonal matrices will appear frequently in Chapter 7.

EXAMPLE 7 The matrix

$$U = \begin{bmatrix} 3/\sqrt{11} & -1/\sqrt{6} & -1/\sqrt{66} \\ 1/\sqrt{11} & 2/\sqrt{6} & -4/\sqrt{66} \\ 1/\sqrt{11} & 1/\sqrt{6} & 7/\sqrt{66} \end{bmatrix}$$

is an orthogonal matrix because it is square and because its columns are orthonormal, by Example 5. Verify that the rows are orthonormal, too!

PRACTICE PROBLEMS

1. Let $\mathbf{u}_1 = \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$ and $\mathbf{u}_2 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$. Show that $\{\mathbf{u}_1, \mathbf{u}_2\}$ is an orthonormal basis for $\mathbb{R}^2$.
2. Let $\mathbf{y}$ and L be as in Example 3 and Fig. 3. Compute the orthogonal projection $\hat{\mathbf{y}}$ of $\mathbf{y}$ onto L using $\mathbf{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ instead of the $\mathbf{u}$ in Example 3.
3. Let U and $\mathbf{x}$ be as in Example 6, and let $\mathbf{y} = \begin{bmatrix} -3\sqrt{2} \\ 6 \end{bmatrix}$. Verify that $U\mathbf{x}\cdot U\mathbf{y} = \mathbf{x}\cdot\mathbf{y}$.

6.2 EXERCISES

In Exercises 1–6, determine which sets of vectors are orthogonal.

1. $\begin{bmatrix} -1 \\ 4 \\ -3 \end{bmatrix}, \begin{bmatrix} 5 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ -4 \\ -7 \end{bmatrix}$

2. $\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -5 \\ -2 \\ 1 \end{bmatrix}$

3. $\begin{bmatrix} 2 \\ -7 \\ -1 \end{bmatrix}, \begin{bmatrix} -6 \\ -3 \\ 9 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix}$

4. $\begin{bmatrix} 2 \\ -5 \\ -3 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ -2 \\ 6 \end{bmatrix}$

5. $\begin{bmatrix} 3 \\ -2 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 3 \\ -3 \\ 4 \end{bmatrix}, \begin{bmatrix} 3 \\ 8 \\ 7 \\ 0 \end{bmatrix}$

6. $\begin{bmatrix} 5 \\ -4 \\ 0 \\ 3 \end{bmatrix}, \begin{bmatrix} -4 \\ 1 \\ -3 \\ 8 \end{bmatrix}, \begin{bmatrix} 3 \\ 3 \\ 5 \\ -1 \end{bmatrix}$

In Exercises 7–10, show that $\{\mathbf{u}_1, \mathbf{u}_2\}$ or $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal basis for $\mathbb{R}^2$ or $\mathbb{R}^3$, respectively. Then express $\mathbf{x}$ as a linear combination of the $\mathbf{u}$'s.

7. $\mathbf{u}_1 = \begin{bmatrix} 2 \\ -3 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$, and $\mathbf{x} = \begin{bmatrix} 9 \\ -7 \end{bmatrix}$

[1] A better name might be *orthonormal matrix*, and this term is found in some statistics texts. However, *orthogonal matrix* is the standard term in linear algebra.

8. $\mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -2 \\ 6 \end{bmatrix}$, and $\mathbf{x} = \begin{bmatrix} -6 \\ 3 \end{bmatrix}$

9. $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -1 \\ 4 \\ 1 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$, and $\mathbf{x} = \begin{bmatrix} 8 \\ -4 \\ -3 \end{bmatrix}$

10. $\mathbf{u}_1 = \begin{bmatrix} 3 \\ -3 \\ 0 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}$, and $\mathbf{x} = \begin{bmatrix} 5 \\ -3 \\ 1 \end{bmatrix}$

11. Compute the orthogonal projection of $\begin{bmatrix} 1 \\ 7 \end{bmatrix}$ onto the line through $\begin{bmatrix} -4 \\ 2 \end{bmatrix}$ and the origin.

12. Compute the orthogonal projection of $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ onto the line through $\begin{bmatrix} -1 \\ 3 \end{bmatrix}$ and the origin.

13. Let $\mathbf{y} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ and $\mathbf{u} = \begin{bmatrix} 4 \\ -7 \end{bmatrix}$. Write $\mathbf{y}$ as the sum of two orthogonal vectors, one in Span $\{\mathbf{u}\}$ and one orthogonal to $\mathbf{u}$.

14. Let $\mathbf{y} = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$ and $\mathbf{u} = \begin{bmatrix} 7 \\ 1 \end{bmatrix}$. Write $\mathbf{y}$ as the sum of a vector in Span $\{\mathbf{u}\}$ and a vector orthogonal to $\mathbf{u}$.

15. Let $\mathbf{y} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ and $\mathbf{u} = \begin{bmatrix} 8 \\ 6 \end{bmatrix}$. Compute the distance from $\mathbf{y}$ to the line through $\mathbf{u}$ and the origin.

16. Let $\mathbf{y} = \begin{bmatrix} -3 \\ 9 \end{bmatrix}$ and $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Compute the distance from $\mathbf{y}$ to the line through $\mathbf{u}$ and the origin.

In Exercises 17–22, determine which sets of vectors are orthonormal. If a set is only orthogonal, normalize the vectors to produce an orthonormal set.

17. $\begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix}$, $\begin{bmatrix} -1/2 \\ 0 \\ 1/2 \end{bmatrix}$

18. $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}$

19. $\begin{bmatrix} -.6 \\ .8 \end{bmatrix}$, $\begin{bmatrix} .8 \\ .6 \end{bmatrix}$

20. $\begin{bmatrix} -2/3 \\ 1/3 \\ 2/3 \end{bmatrix}$, $\begin{bmatrix} 1/3 \\ 2/3 \\ 0 \end{bmatrix}$

21. $\begin{bmatrix} 1/\sqrt{10} \\ 3/\sqrt{20} \\ 3/\sqrt{20} \end{bmatrix}$, $\begin{bmatrix} 3/\sqrt{10} \\ -1/\sqrt{20} \\ -1/\sqrt{20} \end{bmatrix}$, $\begin{bmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$

22. $\begin{bmatrix} 1/\sqrt{18} \\ 4/\sqrt{18} \\ 1/\sqrt{18} \end{bmatrix}$, $\begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix}$, $\begin{bmatrix} -2/3 \\ 1/3 \\ -2/3 \end{bmatrix}$

In Exercises 23 and 24, all vectors are in $\mathbb{R}^n$. Mark each statement True or False. Justify each answer.

23. a. Not every linearly independent set in $\mathbb{R}^n$ is an orthogonal set.

b. If $\mathbf{y}$ is a linear combination of nonzero vectors from an orthogonal set, then the weights in the linear combination can be computed without row operations on a matrix.

c. If the vectors in an orthogonal set of nonzero vectors are normalized, then some of the new vectors may not be orthogonal.

d. A matrix with orthonormal columns is an orthogonal matrix.

e. If L is a line through $\mathbf{0}$ and if $\hat{\mathbf{y}}$ is the orthogonal projection of $\mathbf{y}$ onto L, then $\|\hat{\mathbf{y}}\|$ gives the distance from $\mathbf{y}$ to L.

24. a. Not every orthogonal set in $\mathbb{R}^n$ is linearly independent.

b. If a set $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ has the property that $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ whenever $i \neq j$, then S is an orthonormal set.

c. If the columns of an $m \times n$ matrix A are orthonormal, then the linear mapping $\mathbf{x} \mapsto A\mathbf{x}$ preserves lengths.

d. The orthogonal projection of $\mathbf{y}$ onto $\mathbf{v}$ is the same as the orthogonal projection of $\mathbf{y}$ onto $c\mathbf{v}$ whenever $c \neq 0$.

e. An orthogonal matrix is invertible.

25. Prove Theorem 7. [*Hint:* For (a), compute $\|U\mathbf{x}\|^2$, or prove (b) first.]

26. Suppose W is a subspace of $\mathbb{R}^n$ spanned by n nonzero orthogonal vectors. Explain why $W = \mathbb{R}^n$.

27. Let U be a square matrix with orthonormal columns. Explain why U is invertible. (Mention the theorems you use.)

28. Let U be an $n \times n$ orthogonal matrix. Show that the rows of U form an orthonormal basis of $\mathbb{R}^n$.

29. Let U and V be $n \times n$ orthogonal matrices. Explain why UV is an orthogonal matrix. [That is, explain why UV is invertible and its inverse is $(UV)^T$.]

30. Let U be an orthogonal matrix, and construct V by interchanging some of the columns of U. Explain why V is an orthogonal matrix.

31. Show that the orthogonal projection of a vector $\mathbf{y}$ onto a line L through the origin in $\mathbb{R}^2$ does not depend on the choice of the nonzero $\mathbf{u}$ in L used in the formula for $\hat{\mathbf{y}}$. To do this, suppose $\mathbf{y}$ and $\mathbf{u}$ are given and $\hat{\mathbf{y}}$ has been computed by formula (2) in this section. Replace $\mathbf{u}$ in that formula by $c\mathbf{u}$, where c is an unspecified nonzero scalar. Show that the new formula gives the same $\hat{\mathbf{y}}$.

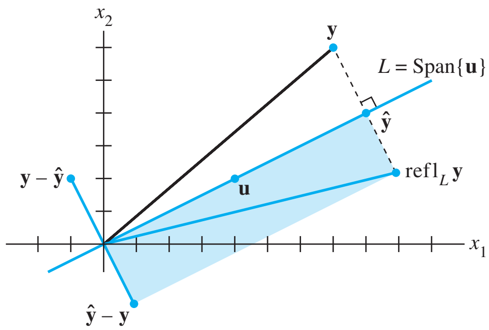
32. Let $\{\mathbf{v}_1, \mathbf{v}_2\}$ be an orthogonal set of nonzero vectors, and let c_1, c_2 be any nonzero scalars. Show that $\{c_1\mathbf{v}_1, c_2\mathbf{v}_2\}$ is also an orthogonal set. Since orthogonality of a set is defined in terms of pairs of vectors, this shows that if the vectors in an orthogonal set are normalized, the new set will still be orthogonal.

33. Given $\mathbf{u} \neq \mathbf{0}$ in $\mathbb{R}^n$, let $L = \text{Span}\{\mathbf{u}\}$. Show that the mapping $\mathbf{x} \mapsto \text{proj}_L\, \mathbf{x}$ is a linear transformation.

34. Given $\mathbf{u} \neq \mathbf{0}$ in $\mathbb{R}^n$, let $L = \text{Span}\{\mathbf{u}\}$. For $\mathbf{y}$ in $\mathbb{R}^n$, the **reflection of y in** L is the point $\text{refl}_L\, \mathbf{y}$ defined by

$\text{refl}_L\, \mathbf{y} = 2 \cdot \text{proj}_L\, \mathbf{y} - \mathbf{y}$

See the figure, which shows that $\text{refl}_L\, \mathbf{y}$ is the sum of $\hat{\mathbf{y}} = \text{proj}_L\, \mathbf{y}$ and $\hat{\mathbf{y}} - \mathbf{y}$. Show that the mapping $\mathbf{y} \mapsto \text{refl}_L\, \mathbf{y}$ is a linear transformation.



The reflection of $\mathbf{y}$ in a line through the origin.

35. [**M**] Show that the columns of the matrix A are orthogonal by making an appropriate matrix calculation. State the calculation you use.

$$A = \begin{bmatrix} -6 & -3 & 6 & 1 \\ -1 & 2 & 1 & -6 \\ 3 & 6 & 3 & -2 \\ 6 & -3 & 6 & -1 \\ 2 & -1 & 2 & 3 \\ -3 & 6 & 3 & 2 \\ -2 & -1 & 2 & -3 \\ 1 & 2 & 1 & 6 \end{bmatrix}$$

36. [**M**] In parts (a)–(d), let U be the matrix formed by normalizing each column of the matrix A in Exercise 35.

a. Compute U^TU and UU^T. How do they differ?

b. Generate a random vector $\mathbf{y}$ in $\mathbb{R}^8$, and compute $\mathbf{p} = UU^T\mathbf{y}$ and $\mathbf{z} = \mathbf{y} - \mathbf{p}$. Explain why $\mathbf{p}$ is in Col A. Verify that $\mathbf{z}$ is orthogonal to $\mathbf{p}$.

c. Verify that $\mathbf{z}$ is orthogonal to each column of U.

d. Notice that $\mathbf{y} = \mathbf{p} + \mathbf{z}$, with $\mathbf{p}$ in Col A. Explain why $\mathbf{z}$ is in $(\text{Col } A)^\perp$. (The significance of this decomposition of $\mathbf{y}$ will be explained in the next section.)

SOLUTIONS TO PRACTICE PROBLEMS

1. The vectors are orthogonal because

$$\mathbf{u}_1 \cdot \mathbf{u}_2 = -2/5 + 2/5 = 0$$

They are unit vectors because

$$\|\mathbf{u}_1\|^2 = (-1/\sqrt{5})^2 + (2/\sqrt{5})^2 = 1/5 + 4/5 = 1$$
$$\|\mathbf{u}_2\|^2 = (2/\sqrt{5})^2 + (1/\sqrt{5})^2 = 4/5 + 1/5 = 1$$

In particular, the set $\{\mathbf{u}_1, \mathbf{u}_2\}$ is linearly independent, and hence is a basis for $\mathbb{R}^2$ since there are two vectors in the set.

2. When $\mathbf{y} = \begin{bmatrix} 7 \\ 6 \end{bmatrix}$ and $\mathbf{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$,

$$\hat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}}\mathbf{u} = \frac{20}{5}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = 4\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

This is the same $\hat{\mathbf{y}}$ found in Example 3. The orthogonal projection does not seem to depend on the $\mathbf{u}$ chosen on the line. See Exercise 31.

3. $$U\mathbf{y} = \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix}\begin{bmatrix} -3\sqrt{2} \\ 6 \end{bmatrix} = \begin{bmatrix} 1 \\ -7 \\ 2 \end{bmatrix}$$

Also, from Example 6, $\mathbf{x} = \begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix}$ and $U\mathbf{x} = \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix}$. Hence

SG Mastering: Orthogonal Basis 6-4

$$U\mathbf{x} \cdot U\mathbf{y} = 3 + 7 + 2 = 12, \quad \text{and} \quad \mathbf{x} \cdot \mathbf{y} = -6 + 18 = 12$$